

The use of individual patient-level data (IPD) to quantify the impact of pretreatment predictors of response to treatment in chronic hepatitis B patients

Shehzad Ali,¹ Stuart Mealing,¹ Neil Hawkins,^{1,2} Benedicte Lescrauwaet,³ Stefan Bjork,⁴ Lorenzo Mantovani,⁵ Pietro Lampertico⁶

To cite: Ali S, Mealing S, Hawkins N, *et al*. The use of individual patient-level data (IPD) to quantify the impact of pretreatment predictors of response to treatment in chronic hepatitis B patients. *BMJ Open* 2013;**3**:e001309. doi:10.1136/bmjopen-2012-001309

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2012-001309>).

Received 23 May 2012
Revised 12 December 2012
Accepted 13 December 2012

This final article is available for use under the terms of the Creative Commons Attribution Non-Commercial 2.0 Licence; see <http://bmjopen.bmj.com>

For numbered affiliations see end of article.

*Correspondence to Stuart Mealing; Stuart.mealing@oxfordoutcomes.com

ABSTRACT

Objectives: Evidence synthesis is an integral part decision-making by reimbursement agencies. When direct evidence is not available, network-meta-analysis (NMA) techniques are commonly used. This approach assumes that the trials are sufficiently similar in terms of treatment-effect modifiers. When imbalances in potential treatment-effect modifiers exist, the NMA approach may not produce fair comparisons. The objective of this study was to identify and quantify the interaction between treatment-effect and potential treatment-effect modifiers, including time-of-response measurement and baseline viral load in chronic hepatitis B (CHB) patients.

Design: Retrospective patient-level data econometric analysis.

Participants: 1353 individuals from two randomised controlled trials of nucleoside-naïve CHB taking 0.5 mg entecavir (n=679) or 100 mg lamivudine (n=668) daily for 48 weeks.

Interventions: Hepatitis B virus (HBV) DNA levels for both drugs were measured at baseline and weeks 24, 36 and 48. Generalised estimating equation for repeated binary responses was used to identify treatment-effect modifiers for response defined at ≤ 400 or ≤ 300 copies/ml.

Primary outcome measures: OR at 48 weeks.

Results: The OR for the time-of-response measurement and treatment-effect interaction term was 1.039 ($p=0.00$) and 1.035 ($p=0.00$) when response was defined at ≤ 400 or ≤ 300 copies/ml, respectively. The baseline HBV DNA and treatment-effect interaction OR was 0.94 ($p=0.047$) and 0.95 ($p=0.096$), respectively, for the two response definitions suggesting evidence of interaction between baseline disease activity and treatment effect. The interaction between HBeAg status and treatment effect was not statistically significant.

Conclusions: The measurement time point seems to modify the relative treatment effect of entecavir compared to lamivudine, measured on the OR scale. Evidence also suggested that differences in baseline viral load may also alter relative treatment effect. Meta-analyses should account for such modifiers when generating relative efficacy estimates.

ARTICLE SUMMARY

Article focus

- In patients with chronic hepatitis B, are there any baseline patient characteristics that impact on the probability of achieving undetectable viral load after 1 year of therapy?
- If so, what is the magnitude of the effect?

Key messages

- Time on treatment and baseline viral load were independent treatment-effect modifiers.
- The treatment-effect interaction OR for baseline viral load was 0.94 ($p=0.047$).
- Future meta-analyses should account for this interaction term in order to generate 'like-for-like' comparisons.

Strengths and limitations of this study

- Analyses based on patient-level data from two high-quality randomised controlled trials (patient count=1353).
- The statistical models took account of the longitudinal nature of the data and the correlation between repeated measures.
- HBV DNA levels were observed at relatively small number of discrete time points, that is, 24, 36 and 48 weeks
- Only two treatments included in the trials (entecavir and lamivudine).

INTRODUCTION

Chronic hepatitis B (CHB) is an infectious condition caused by the hepatitis B virus (HBV). Long-term consequences of infection include cirrhosis, hepatocellular carcinoma and ultimately death.¹ For patients in the UK, the licensed treatments for CHB are entecavir (ETV), lamivudine (LAM), telbivudine, adefovir dipivoxil, tenofovir disoproxil fumarate (TDF), peginterferon α -2a, interferon α -2a and interferon α -2b.² These treatments have been evaluated in randomised controlled trials (RCTs); however, head-to-head comparisons

are not always available for all possible treatment comparisons. From the perspective of a reimbursement agency such as the National Institute for Health and Clinical Excellence (NICE), this poses a challenge since relative estimates of the treatment efficacy may not be available for each intervention to inform reimbursement decisions.

One strategy to overcoming this challenge is to use network meta-analysis (NMA) to generate relative efficacy estimates between competing treatments.^{3–6} While a traditional meta-analysis includes studies that compare two interventions with each other, an NMA synthesises all available evidence from a network of RCTs involving multiple treatments compared directly or indirectly or both. NMA extends the concept of a traditional meta-analysis by including multiple pairwise comparisons across a range of interventions and provides estimates of the relative treatment effect on multiple treatment comparisons by using a chain of evidence that links treatments of interest. It represents a comprehensive analysis framework for combining both direct and indirect evidence in such a manner as to preserve randomisation within trials.

The NMA approach allows consideration of all relevant evidence and addresses research questions in the absence of direct comparative evidence, thus improving the precision of estimates by combining direct and indirect evidence. One of the key assumptions underpinning this method is that the studies included in the analysis are homogeneous, that is, the trials are sufficiently similar on study and patient characteristics ('covariates'). If these covariates act as modifiers of the relative treatment effect (ie, the difference in response on a given scale, typically log-odds or odds) and their distribution is not balanced across the studies that are compared directly or indirectly, the similarity assumption is violated and the NMA is affected by confounding bias. Effect modifiers may include patient characteristics, study setting, length of follow-up, outcome definition and measurement and study methodology (eg, protocol requirements and study time frame).⁷ This is particularly important in cases where response to treatment is defined in terms of post-treatment level of a measure (eg, a biological marker such as response viral load) when the baseline level of this measure is known to vary across studies. As a result, if one study happens to recruit patients with a more advanced disease (ie, worse levels of a given clinical variable), and this variable is known to modify the impact of treatment, then the level of response achieved in this study is likely to be smaller compared to another study which primarily includes patients with less-advanced disease (ie, with better baseline variable levels), all things being equal. To identify treatment-effect modifiers and to provide an estimate of the magnitude of treatment-covariate interaction, regression analysis of individual patient-level data can be used. This, in turn, can form the basis for evaluating baseline imbalances across trials included in the NMA. It should, however, be noted that such within-trial

interaction analysis would provide an estimate for the magnitude of potential treatment modifier effect rather than proving beyond doubt that these interactions exist at the study level.

To date, two network meta-analyses of treatments for CHB have been published in peer-review journals.^{8–9} Both studies defined response to treatment in terms of attainment of undetectable levels of HBV DNA below a predefined threshold at a given time point. However, in both cases, the authors did not attempt to fully account for the differences in different sources of heterogeneity, including baseline patient characteristics across different trials. Moreover, the effect of using different definitions of virological response to treatment was not explored during evidence synthesis.

The objective of this study was to explore and quantify the relationship between treatment effect and patient characteristics, in particular, baseline disease severity and time-of-response measurement, in predicting response to CHB treatment. Baseline HBV DNA level as measured using the PCR assay is known to be a predictor of treatment response but the extent of this relationship, and whether or not it is also a treatment-effect modifier (ie, whether or not it impacts on the efficacy of a particular treatment option) is not currently known. This work is of interest to the clinical community since it would provide evidence on whether and how much treatment effects vary based on these variables. It will also be of value in future network meta-analyses since it explores the methodological implications of treatment interactions that can impact on the observed level of response. Extending network meta-analysis models with treatment-by-covariate interactions may explain heterogeneity in relative treatment effects.

We used individual patient level data (IPD) from two clinical trials of CHB patients, comparing ETV and LAM, to quantify the relationship between treatment effect and baseline covariates and time-of-response measurement to identify any treatment-effect modifiers. Treatment response was defined in terms of attainment of undetectable levels of HBV DNA below predefined threshold values of ≤ 400 and ≤ 300 copies/ml.

METHODS

Patient-level data from 1353 treatment-naïve individuals recruited into two multinational double-blinded, double-dummy studies of ETV 0.5 mg once daily (n=679) and LAM 100 mg once daily (n=668) were made available by Bristol-Myers Squibb for the purpose of this analysis (study names: 022 and 027).^{10–11} In both clinical trials, all individuals were nucleoside-naïve with the common composite primary efficacy endpoint being histological improvement (\geq two point decrease in the Knodell necroinflammatory score) and no worsening of fibrosis (\geq one point increase in the Knodell fibrosis score), at week 48, compared to baseline. Individuals were

HBeAg-positive in Study 022 and HBeAg-negative in the Study 027.

The primary rationale for the study was to explore the effect of interaction between baseline variables and treatment effect in predicting response. Here, response is defined as a binary (yes/no) variable in terms of achieving a threshold of undetectable levels of HBV DNA. Two commonly used response thresholds were evaluated in this study, that is, ≤ 400 or ≤ 300 copies/ml measured by PCR assay. Response was evaluated for each individual patient at weeks 24, 36 and 48. The following covariates were used in the statistical analysis: natural log of baseline HBV DNA measured by PCR assay (variable: 'LPCR_0'), treatment received ('ETV' or 'LAM', with LAM used as the reference group), HBeAg antigen status of the patient ('HBeAg') and a time variable ('TIME') expressed in weeks to evaluate the impact of duration on treatment response. The following interaction terms were included: treatment with ETV and baseline viral load ('ETV*PCR'), treatment with ETV and time ('ETV*TIME') and treatment with ETV and e-antigen status (ETV*HBeAg). Study-specific intercepts were also included in the analysis. Covariates were centred at their mean values, except for the time variable which was centred at 48 weeks.

Statistical analysis was performed using the generalised estimating equations (GEE) with logit link function and autoregressive correlation structure. The use of GEE is preferred when outcomes are correlated and the focus of analysis is on estimating average effect in the population. The GEE takes an account of the correlation between repeated observations from the same individual at multiple time points. Moreover, since the sample size needed to achieve adequate statistical power to detect interaction effects is larger than the sample size needed to detect main effects,¹² the use of GEE is preferred over cross-sectional logistic regression as it allows the use of multiple observations per individual. The analysis was conducted in Stata V.11 (StataCorp. 2009. Stata Statistical Software: Release 11. College Station, Texas: StataCorp LP).

RESULTS

To contextualise the results of the statistical analysis conducted in this study, we present the summary statistics from a systematic review of CHB interventions conducted separately by the authors and presented at the International Society for Pharmacoeconomics and Outcomes Research conference in 2011. Table 1 presents baseline HBV DNA levels in the RCTs of CHB interventions. The table shows that the baseline HBV DNA values in all trials ranged from 5.6 to 10.3 log₁₀ copies/ml.

For HBeAg-positive patients, the range was between 5.7 and 10.3 log₁₀ copies/ml, whereas for HBeAg-negative patients, the range was 5.6–7.8 log₁₀ copies/ml. Of note, however, is that when the focus is

restricted to the key HBeAg-positive regulatory trials for ETV and TDF (AI463-022 and Marcellin2008), there is an approximate difference of 1 log₁₀ copies/ml. The corresponding difference between the values in the ETV and TDF in HBeAg-negative regulatory studies (AI463-023 and Marcellin2008) was approximately 0.5 log₁₀ copies/ml. Hence, the studies appear dissimilar in terms of baseline viral load levels.

GEE logit regression models evaluated the odds of achieving treatment response at two threshold values (ie, HBV DNA ≤ 400 or ≤ 300 copies/ml) after adjusting for baseline characteristics. Both main effect and interaction terms were included in the analysis. The results are presented as both log-ORs (tables 2 and 3) and ORs (tables 4 and 5) with CIs. An OR of <1 suggests a decrease in odds of achieving treatment response when the level of baseline variable increases by one unit, and vice versa.

The results show that the coefficient on the interaction term for treatment and baseline log-PCR (ETV*PCR) is negative for both threshold values (tables 2 and 3) with different levels of statistical significance. The odds for the interaction term for response at ≤ 400 copies/ml were found to be 0.94 which represents the multiplicative factor by which the ratio of the predicted odds of response for ETV and predicted odds for LAM changes when the baseline log-PCR increases by one unit. In other words, when the baseline log-PCR increases by one unit from its mean value, the OR of response for ETV compared to LAM decreases by 5.6% (ie, $(1-0.94)*100$). The analysis at ≤ 300 copies/ml found that the odds were similar, that is, 0.95. However, it should be noted that the coefficients are only marginally significant in both analyses ($p=0.047$ for ≤ 400 copies/ml and $p=0.096$ for ≤ 300 copies/ml). These results show that there is evidence (albeit relatively weak) to suggest that the treatment effect may be moderated by the level of baseline log-PCR.

The interaction analysis of treatment and TIME in the GEE model is also statistically significant ($p=0.00$) for both response definitions. This indicates that the treatment effect is moderated by time-of-response measurement. The positive coefficient on the interaction variable suggests that when time-of-response measurement increases by 1 week, the OR of response for ETV compared to LAM increases by 3.9%. This indicates that, while the odds of achieving undetectable levels of HBV DNA increases with time in both treatment groups, the rate of change in odds is higher for ETV compared to LAM. We also evaluated the interaction between HBeAg status and treatment effect which was found to be statistically non-significant in both analyses. This result shows that there is no evidence to suggest that the odds of response for ETV compared to LAM is moderated by the HBeAg status.

CONCLUSION

This study evaluated the effect of interaction between treatment received and the baseline characteristics and

Table 1 Baseline HBV DNA levels by study

Study	Treatment	Baseline HBV DNA levels (log ₁₀ copies/ml)	
		HBeAg Positive	HBeAg Negative
018 Study Group	TBV 600 mg	9.57	
	ADV 10 mg	9.98	
ADV 437 Study Group	Placebo	8.12	
	ADV 10 mg	8.25	
Study AI463-022	ETV 0.5 mg	9.62	
	LAM 100 mg	9.69	
Cooksley 2003	IFNA	9.29	
	PegIFNA	9.25	
Hou 2008	TBV 600 mg	9.3	
	LAM 100 mg	9.7	
Lau 2005	PegIFNA	9.9	
	LMV 100 mg	10.1	
Leung 2009	ETV 0.5 mg	10.3	
	ADV 10 mg	9.88	
Marcellin 2008	TDF 300 mg	8.64	
	ADV 10 mg	8.88	
Ren 2007	LAM 100 mg	8.49	
	ETV 0.5 mg	8.52	
Shindo 2009	LAM 100 mg	7.94	
	ETV 0.5 mg	8.39	
USLIG	Placebo	5.7	
	LAM 100 mg	10.2	
	Placebo	8.6	
AI463023	ETV 0.5 mg	8.8	7.7
	LAM 100 mg	8.7	7.6
Globe study group	LAM 100 mg	9.5	7.4
	TBV 600 mg	9.5	7.7
Study AI463-027	ETV 0.5 mg		7.6
	LAM 100 mg		7.6
Chan 2007	Placebo		5.6
	LAM100 mg		5.7
Hou 2008	TBV 600 mg		7.8
	LAM 100 mg		7.6
Marcellin 2004	PegIFNA		7.1
	LAM 100 mg		7.2
Marcellin 2008	TDF 300 mg		6.9
	ADV 10 mg		7.0

ADV, adefovir dipivoxil; ETV, entecavir; HBV, hepatitis B virus; LAM, lamivudine ; TBV, telbivudine; TDF, tenofovir disoproxil fumarate.

Table 2 Log-odds of response based on generalised estimating equation for treatment response at ≤400 copies/ml

Variables	Log-odds	SE	z	P> z	(95% CI)
Constant	1.618	1.118	1.450	0.148	−0.573 to 3.809
Main effects					
LPCR_0 (centred)	−0.172	0.020	−8.490	0.000	−0.212 to −0.132
Treatment, ETV	1.563	0.224	6.990	0.000	1.124 to 2.001
HBeAg (positive)	−1.277	0.567	−2.250	0.024	−2.387 to −0.166
Time (centred at 48 weeks)	0.022	0.005	4.480	0.000	0.012 to 0.031
Clinical study (ID=27)/	−0.306	0.558	−0.550	0.583	−1.399 to 0.787
Interaction terms					
ETV * time (centred at 48 weeks)	0.039	0.008	5.040	0.000	0.024 to 0.054
ETV *PCR (centred at 48 weeks)	−0.059	0.030	−1.980	0.047	−0.117 to −0.001
ETV * HBeAg (positive)	0.267	0.249	1.070	0.284	−0.221 to 0.754

ETV, entecavir.

Table 3 Log-odds of response based on generalised estimating equation for treatment response at ≤ 300 copies/ml

Variables	Log-odds	SE	z	P> z	(95% CI)
Constant	1.817	1.128	1.610	0.107	−0.394 to 4.027
Main effects					
LPCR_0 (centred)	−0.166	0.020	−8.160	0.000	−0.205 to −0.126
Treatment, ETV	1.429	0.214	6.680	0.000	1.010 to 1.848
HBeAg (positive)	−1.414	0.572	−2.470	0.013	−2.536 to −0.293
Time (centred at 48 weeks)	0.025	0.005	5.290	0.000	0.016 to 0.035
Clinical study (ID=27)	−0.450	0.563	−0.800	0.424	−1.553 to 0.653
Interaction terms					
ETV * Time (centered at 48 weeks)	0.035	0.007	4.710	0.000	0.020 to 0.049
ETV *PCR (centered at 48 weeks)	−0.049	0.029	−1.660	0.096	−0.106 to 0.009
ETV * HBeAg (positive)	0.264	0.244	1.080	0.278	−0.213 to 0.742

ETV, entecavir; LPCR_0, natural log of baseline HBV DNA measured by PCR assay.

time-of-response measurement in predicting the odds of treatment response (defined as undetectable HBV DNA levels at threshold values of 400 and 300 copies/ml). The baseline variables included the log of baseline HBV DNA level, the time-of-response measurement (in weeks) and HBeAg status. While the qualitative relationship between time, baseline viral load and response has been documented in the literature,¹³ the key aim of the analysis was to identify and quantify the interaction effects that may act as treatment-effect modifiers which, in turn, may be useful to adjust baseline differences in future meta-analyses.

The analyses presented in this paper showed that there is strong evidence to suggest that time-of-response measurement may act as treatment-effect modifier. This suggests that time of measurement should be taken into account during NMA, when clinical trials with different periods of patient follow-up are included in the analysis. Our analysis also found that there is weak evidence to suggest that the baseline HBV DNA level may also interact with treatment effect. This is potentially an important finding to suggest that future network meta-analyses should evaluate the need to account for differences in baseline disease activity in CHB patients. We also evaluated the main (non-interactive) impact of baseline HBV

DNA level on odds of response in the models. The coefficient suggests that patients with more severe or advanced levels of disease activity before treatment initiation (as measured by baseline viral load) are less likely to achieve response (as measured by response viral load) at any given time point, that is, higher baseline disease activity predicts worse response. This is in line with earlier findings.

We did not find any significant interaction between treatment-effect and HBeAg status. However, HBeAg status was found to be a significant predictor as a main effect suggesting that patients with positive HBeAg status are less likely to achieve response after controlling for baseline HBV DNA and time-of-response measurement.

Reimbursement agencies, such as NICE, typically make decisions based on evidence regarding relative treatment effects based on clinical efficacy data on end-points, such as the level of response achieved after treatment for a certain period of time, as a surrogate for clinical efficacy.¹⁴ From the perspective of a reimbursement agency, the key finding of this analysis is that the time-of-response measurement and baseline disease activity may act as treatment-effect modifiers, suggesting that when these variables are not distributed in a balanced way across trials, there is potential for

Table 4 Odds of response based on generalised estimating equation for treatment response at ≤ 400 copies/ml

Variables	Odds	SE	z	P> z	(95% CI)
Main effects					
LPCR_0 (centred)	0.842	0.017	−8.490	0.000	0.809 to 0.876
Treatment, ETV	4.773	1.068	6.990	0.000	3.079 to 7.400
HBeAg (positive)	0.279	0.158	−2.250	0.024	0.092 to 0.847
Time (centred at 48 weeks)	1.022	0.005	4.480	0.000	1.012 to 1.032
Clinical study (ID=27)	0.736	0.411	−0.550	0.583	0.247 to 2.196
Interaction terms					
ETV * time (centred at 48 weeks)	1.039	0.008	5.040	0.000	1.024 to 1.055
ETV *PCR (centred at 48 weeks)	0.943	0.028	−1.980	0.047	0.889 to 0.999
ETV * HBeAg (positive)	1.305	0.325	1.070	0.284	0.801 to 2.126

ETV, entecavir; LPCR_0, natural log of baseline HBV DNA measured by PCR assay.

Table 5 Odds of response based on generalised estimating equation for treatment response at ≤ 300 copies/ml

Variables	Odds	SE	z	P> z	(95% CI)
Main effects					
Main effects	0.847	0.017	-8.160	0.000	0.814 to 0.882
LPCR_0 (centred)	4.175	0.893	6.680	0.000	2.745 to 6.349
Treatment, ETV	0.243	0.139	-2.470	0.013	0.079 to 0.746
HBeAg (positive)	1.026	0.005	5.290	0.000	1.016 to 1.035
Time (centred at 48 weeks)	0.637	0.359	-0.800	0.424	0.212 to 1.921
Clinical study (ID=27)					
Interaction terms	1.035	0.008	4.710	0.000	1.020 to 1.050
ETV * Time (centered at 48 wks)	0.952	0.028	-1.660	0.096	0.899 to 1.009
ETV *PCR (centered at 48 wks)	1.303	1.303	1.080	0.278	0.808 to 2.101

ETV, entecavir; LPCR_0, natural log of baseline HBV DNA measured by PCR assay.

confounding bias in the resulting meta-analysis estimates. Hence, analysts should evaluate differences in patient-level and study-level characteristics that may act as treatment-effect modifiers while undertaking using direct, indirect and mixed treatment comparisons to allow fair 'like-for-like' comparisons to be made. However, it should be noted that the within-trial interaction analysis provides an estimate of the magnitude of potential treatment-effect modifier rather than proving beyond doubt that these interactions exist at the study level.

All mixed treatment comparisons of CHB treatments published to date did not adjust for patient baseline characteristics. Therefore, if heterogeneity or treatment modifiers across clinical trials are not accounted for, then meta-analysis may produce biased estimates in favour of treatment(s) that had patients with relatively longer time-of-response measurement and less severe disease activity in the trials (ie, lower baseline risk). Furthermore, when such evidence is incorporated in an economic model, such as those developed by Dakin *et al*¹⁵ or Veenstra *et al*,^{16 17} the cost-effectiveness estimates may be biased against interventions that was studied in patients with more severely disease activity.

Strengths and weaknesses of the analyses

The study used individual patient level data from two longitudinal randomised trials over a period of 48 weeks. The statistical models took account of the longitudinal nature of the data and the correlation between repeated measures. Interaction effects that may act as treatment-effect modifiers were evaluated and quantified in this study.

One potential limitation of the study is that response HBV DNA levels were observed at relatively small number of discrete time points, that is, 24, 36 and 48 weeks. This may have an implication on the statistical power of the study to detect significant effects. Another limitation is that the relationship between treatment effect and patient-level baseline characteristics was explored only in the trials of ETV versus LAM in this study, and as such, we were not able to explore treatment-effect modifiers in other interventions such as adefovir and tenofovir. It is difficult to anticipate whether or not such interaction

effects will be observed in other CHB studies. However, we do not know of any reasons why the observed relationship may not hold for CHB patients receiving other treatments. However, we would recommend that such analyses are repeated in other randomised trials to ascertain the validity of these findings. Finally, the analysis assumed that the unobserved patient characteristics in this study did not directly influence the observed relationship between treatment-effect and baseline characteristics.

Implications for evidence synthesis

This study found that the time-of-response measurement and the level of baseline disease activity may act as treatment-effect modifiers in CHB patients. From the perspective of evidence synthesis, this study identifies an important issue, that is, treatment-effect modifiers may impact on the comparisons made across several studies in meta-analyses. It highlights the need to explore the impact of baseline characteristics and time-period imbalances between studies included in meta-analyses. When effect-modifier relationships are significant, comparisons across treatments may not be fair and could potentially bias the comparative treatment-effect estimates.

Author affiliations

¹Oxford Outcomes Ltd, Oxford, UK

²Centre for Health Economics, University of York, York, UK

³Xintera Consulting Bvba, Leuven, Belgium

⁴Bristol-Myers Squibb Sarl, Rueil Malmaison, Paris, France

⁵University of Naples, Naples, Italy

⁶First Division of Gastroenterology, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Università degli Studi di Milano, Milano, Italy

Contributors SA performed all analyses and contributed to the preparation of the manuscript. SM and NH assisted SA in performing the analysis and assisted in the preparation of the manuscript. BL, SB, LM and PL assisted in identifying key parameters, validated all results and assisted in the preparation of the manuscript.

Funding This study was funded by Bristol-Myers Squibb.

Competing interests SA, SM and NH worked for Oxford Outcomes Ltd who received funding from Bristol-Myers Squibb to conduct this analysis. BL and SB were Bristol-Myers Squibb employees during the conduct of the study, and received funding for consulting services on specific HEOR projects. PL received honoraria for advisory boards and speaking bureau for Bristol-Myers

Squibb. LM received honoraria from Bristol-Myers Squibb for advisory boards, but not in the field of hepatology nor in one of the infectiology.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. European Association for the Study of the Liver. EASL clinical practice guidelines: management of chronic hepatitis B. *J Hepatol* 2009;50:227–42.
2. British National Formulary. <http://www.bnf.org/bnf/> (accessed 16 Jan 2012).
3. Caldwell D, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2007;331:897–900.
4. Jansen JP, Crawford B, Bergman G, *et al.* Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value Health* 2008;11:956–64.
5. Ades AE, Sculpher M, Sutton A, *et al.* Bayesian methods for evidence synthesis in cost-effectiveness analysis. *Pharmacoeconomics* 2006;24:1–19.
6. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105–24.
7. Song F, Loke YK, Walsh T, *et al.* Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009;338:b1147.
8. Woo G, Tomlinson G, Nishikawa Y, *et al.* Tenofovir and entecavir are the most effective antiviral agents for chronic hepatitis B: a systematic review and bayesian meta-analysis. *Gastroenterology* 2010;139:1218–29.
9. Dakin HA, Fidler C, Harper CH. Mixed treatment comparison meta-analysis evaluating the relative efficacy of nucleos(t)ides for treatment of nucleos(t)ide-naïve patients with chronic hepatitis B. *Value Health* 2010;13:934–45.
10. Lai CL, Shouval D, Lok AS, *et al.* Entecavir versus lamivudine for patients with HBeAg-negative chronic hepatitis B. *N Engl J Med* 2006;354:1011–20.
11. Chang TT, Gish RG, de MR, *et al.* A comparison of entecavir and lamivudine for HBeAg-positive chronic hepatitis B. *N Engl J Med* 2006;354:1001–10.
12. Maxwell SE, Delaney HD. *Designing experiments and analyzing data: a model comparison perspective*. Taylor & Francis Group, Lawrence Erlbaum, 2004.
13. Colonno RJ, Rose FE, Baldick CJ, *et al.* High barrier to resistance results in no emergence of entecavir resistance in nucleoside-naïve subjects during the first two years of therapy. *J Hepatol* 2006;44 (Suppl 2):S182.
14. National Institute for Health and Clinical Excellence. Final appraisal determination: Entecavir for the treatment of chronic hepatitis B. <http://www.nice.org.uk/nicemedia/pdf/EntecavirFinalFAD.pdf> (accessed 16 Jan 2012).
15. Dakin HA, Bentley A, Dusheiko GM. Cost-utility analysis of tenofovir disoproxil fumarate in the treatment of chronic hepatitis B. *Value Health* 2010;13:922–33.
16. Veenstra DL, Sullivan SD, Clarke L, *et al.* Cost-effectiveness of entecavir versus lamivudine with adefovir salvage in HBeAg-positive chronic hepatitis B based on a 2-year clinical trial. *Pharmacoeconomics* 2007;25:963–77.
17. Veenstra DL, Sullivan SD, Dusheiko GM, *et al.* Cost-effectiveness of peginterferon alpha-2a compared with lamivudine treatment in patients with HBe-antigen-positive chronic hepatitis B in the UK. *Eur J Gastroenterol Hepatol* 2007;19:631–38.